

Nonlinear Eigenvalue Problems and Contour Integrals

Marc Van Barel^{a,1,*}, Peter Kravanja^a

^a*KU Leuven, Department of Computer Science,
Celestijnenlaan 200A, B-3001 Leuven (Heverlee), Belgium*

Abstract

In this paper Beyn's algorithm for solving nonlinear eigenvalue problems is given a new interpretation and a variant is designed in which the required information is extracted via the canonical polyadic decomposition of a Hankel tensor. A numerical example shows that the choice of the filter function is very important, particularly with respect to where it is positioned in the complex plane.

Keywords: nonlinear eigenvalue problems, contour integrals, Keldysh' theorem, Beyn's algorithm, canonical polyadic decomposition, filter function, gun problem

2010 MSC: 65-F15, 47-J10, 30-E05

1. Introduction

In this paper, we consider the following *nonlinear eigenvalue problem*. Given an integer $m \geq 1$, a domain $\Omega \subset \mathbb{C}$ and a matrix-valued function $T : \Omega \rightarrow \mathbb{C}^{m \times m}$ analytic in Ω , we want to determine the values $\lambda \in \Omega$ (eigenvalues) and $v \in \mathbb{C}^m$, $v \neq 0$ (eigenvectors) such that

$$T(\lambda)v = 0.$$

*Corresponding author

Email addresses: `Marc.VanBarel@cs.kuleuven.be` (Marc Van Barel),
`peterkravanja@gmail.com` (Peter Kravanja)

¹The research was partially supported by the Research Council KU Leuven, project OT/10/038 (Multi-parameter model order reduction and its applications), PF/10/002 Optimization in Engineering Centre (OPTEC), by the Fund for Scientific Research–Flanders (Belgium), G.0828.14N (Multivariate polynomial and rational interpolation and approximation), and by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office, Belgian Network DYSCO (Dynamical Systems, Control, and Optimization). The scientific responsibility rests with its author(s).

More specifically, given a closed contour $\Gamma \subset \Omega$ that has its interior in Ω , we consider the problem of approximating *all* eigenvalues (and corresponding eigenvectors) *inside* Γ . We observe that if the problem size m is equal to 1, then
5 the problem reduces to that of computing all the zeros λ of the analytic scalar function T inside the closed contour Γ .

The approach discussed in this paper is based on (numerical approximations of) contour integrals of the resolvent operator $T(z)^{-1}$ applied to a rectangular matrix \hat{V} :

$$\frac{1}{2\pi i} \int_{\Gamma} f(z) T(z)^{-1} \hat{V} dz \in \mathbb{C}^{m \times q}$$

where $f : \Omega \rightarrow \mathbb{C}$ is analytic in Ω and $\hat{V} \in \mathbb{C}^{m \times q}$ is a matrix chosen randomly or in another specified way, with $q \leq m$.

Using contour integrals to solve linear and nonlinear eigenvalue problems is a relatively recent development compared to the history of applying such methods to look for the zeros of a scalar analytic function, a problem that we investigated in [1]. Our approach was based on the pioneering work of Delves and Lyness [2] (see also [3] for a recent overview of numerical algorithms based on analytic function values at roots of unity). We reduced the problem to a generalized eigenvalue problem involving a Hankel matrix as well as a shifted Hankel matrix consisting of the moments of the analytic function T . For generalized linear eigenvalue problems corresponding to the pencil $A - zB$ Tetsuya Sakurai and his co-authors [4, 5, 6, 7] (see also [8] for the specific case where $A, B \in \mathbb{R}^{m \times m}$ are symmetric and B is positive definite) use

$$\frac{1}{2\pi i} \int_{\Gamma} (z - \gamma)^p \hat{u}^H (zB - A)^{-1} \hat{v} dz, \quad p = 0, 1, 2, \dots$$

where $\gamma \in \mathbb{C}$ belongs to the interior of Γ and the vectors $\hat{u}, \hat{v} \in \mathbb{C}^m$ have
10 been chosen at random. Note that these contour integrals are scalar moments based on the resolvent $(zB - A)^{-1}$. Given an upper estimate q for the number of eigenvalues of $A - zB$ located inside Γ , these contour integrals are approximated via a quadrature formula (e.g., the trapezoidal rule if Γ is the unit circle) for $p = 0, 1, \dots, 2q - 1$. The generalized eigenvalue problem (of size $q \times q$) involving

15 the Hankel matrix and the shifted Hankel matrix based on the moments leads to approximations of the eigenvalues of $A - zB$ located inside Γ .

To approximate the eigenvectors, specific linear combinations are taken from the columns of the rectangular matrix given by

$$\frac{1}{2\pi i} \int_{\Gamma} (z - \gamma)^p (zB - A)^{-1} \hat{v} dz \in \mathbb{C}^m$$

for $p = 0, 1, \dots, q - 1$.

Eric Polizzi [9] studied the pencil $A - zB$ where $A, B \in \mathbb{C}^{m \times m}$ are Hermitian and B is positive definite. To compute all the eigenvalues located inside a compact interval on the real axis (enclosed by the contour Γ), he considered

$$S = \frac{1}{2\pi i} \int_{\Gamma} (zB - A)^{-1} B \hat{V} dz$$

for a rectangular matrix $\hat{V} \in \mathbb{C}^{m \times q}$ chosen at random, given an upper estimate q for the number of eigenvalues.

20 Polizzi's FEAST algorithm can be summarized as follows (see also Krämer et al. [10]; Tang and Polizzi [11]; Laux [12]; Güttel et al. [13]):

1. Choose $\hat{V} \in \mathbb{C}^{m \times q}$ of rank q .
2. Compute S by contour integration.
3. Orthogonalize S resulting in the matrix Q having orthonormal columns.
4. Form the Rayleigh quotients

$$A_Q = Q^H A Q \quad \text{and} \quad B_Q = Q^H B Q.$$

5. Solve the size- q generalized eigenvalue problem

$$A_Q \tilde{Y} = B_Q \tilde{Y} \tilde{\Lambda}.$$

- 25 6. Compute the approximate Ritz pairs $(\tilde{\Lambda}, \tilde{X} = Q \tilde{Y})$.
7. If convergence is not reached, then go to Step 1, with $\hat{V} = \tilde{X}$.

To compute all the eigenvalues inside the contour Γ for a nonlinear analytic function T , Tetsuya Sakurai and his co-authors [14, 15] (see also [16] for the specific case of polynomial eigenvalue problems) use the scalars

$$\frac{1}{2\pi i} \int_{\Gamma} z^p \hat{u}^H T(z)^{-1} \hat{v} dz, \quad p = 0, 1, 2, \dots$$

where the vectors $\hat{u}, \hat{v} \in \mathbb{C}^m$ have been chosen at random. To approximate the eigenvectors, they consider the vectors

$$\frac{1}{2\pi i} \int_{\Gamma} z^p T(z)^{-1} \hat{v} dz, \quad p = 0, 1, 2, \dots$$

The eigenvectors are specific linear combinations of these vectors.

In this paper, we present a new interpretation of Beyn's algorithm [17]. Contrary to Beyn, who is not interested in the eigenvalues located outside the
 30 contour Γ , we will indicate that these eigenvalues, as well as the behaviour of the analytic function T outside but near Γ , play an important role to assess the accuracy of the computed eigenvalues. To simplify the technical details, we consider only simple eigenvalues. It is straightforward, however, to extend our approach to multiple eigenvalues by using the results that Beyn has described
 35 for this case.

Our paper is organized as follows. In Section 2, we start by summarizing Beyn's algorithm, which is based on Keldysh' theorem. For the sake of simplicity, we consider only simple eigenvalues. Section 3 describes the effect of approximating the contour integrals by numerical integration and introduces
 40 the corresponding filter function, which we use to reinterpret Beyn's algorithm in Section 4. In Section 5, we propose a variant of the most important substep of Beyn's algorithm: the extraction of the eigenvalue and eigenvector information from the computed moments. Our variant is based on the canonical polyadic decomposition of a Hankel tensor composed from the moments. Section 6 provides
 45 three strategies for solving a specific nonlinear eigenvalue problem. Finally, the conclusions are given in Section 7.

2. Beyn's algorithm

We start by summarizing Beyn's algorithm [17]. His approach is based on Keldysh' theorem, which we recall here for the case of simple eigenvalues.

50 **Theorem 1 (Keldysh [18, 19]).** *Let $\mathcal{C} \subset \Omega$ be a compact subset, let T be a matrix-valued function $T : \Omega \rightarrow \mathbb{C}^{m \times m}$ analytic in Ω and let $n(\mathcal{C})$ denote the number of eigenvalues of T in \mathcal{C} .*

Let λ_k for $k = 1, \dots, n(\mathcal{C})$ denote these eigenvalues and suppose that all of them are simple. Let v_k and w_k for $k = 1, \dots, n(\mathcal{C})$ denote their left and right eigenvectors, such that

$$T(\lambda_k)v_k = 0 \quad w_k^H T(\lambda_k) = 0 \quad w_k^H T'(\lambda_k)v_k = 1.$$

Then there is a neighbourhood \mathcal{U} of \mathcal{C} in Ω and an analytic function $R : \mathcal{U} \rightarrow \mathbb{C}^{m \times m}$ such that the resolvent $T(z)^{-1}$ can be written as

$$T(z)^{-1} = \sum_{k=1}^{n(\mathcal{C})} v_k w_k^H (z - \lambda_k)^{-1} + R(z) \quad (1)$$

for all $z \in \mathcal{U} \setminus \{\lambda_1, \dots, \lambda_{n(\mathcal{C})}\}$.

We observe that if T is a matrix-valued strictly proper rational function, the
 55 analytic function R is equal to zero. This is the case, for example, if $T(z) = A - zB$ with B nonsingular or if $T(z)$ is a matrix polynomial in z with nonsingular highest degree coefficient.

Using expression (1) for the resolvent, we derive the following expression for the corresponding contour integral.

60 **Corollary 1.** *Suppose that T has no eigenvalues on the contour $\Gamma \subset \Omega$ and let $n(\Gamma)$ denote the number of eigenvalues of T inside Γ .*

Let λ_k for $k = 1, \dots, n(\Gamma)$ denote these eigenvalues and suppose that all of them are simple. Let v_k and w_k for $k = 1, \dots, n(\Gamma)$ denote the corresponding left and right (normalized) eigenvectors.

Then

$$\frac{1}{2\pi i} \int_{\Gamma} f(z) T(z)^{-1} dz = \sum_{k=1}^{n(\Gamma)} f(\lambda_k) v_k w_k^H$$

65 *for any function $f : \Omega \rightarrow \mathbb{C}$ that is analytic in Ω .*

With respect to the question of how to develop methods for approximating the eigenvalues inside the contour Γ , the previous corollary informs us that, when “measuring” the resolvent by contour integration with respect to an analytic function f , the result is a sum of rank one terms where only the magnitude of
 70 these rank one terms depends on f .

This observation leads to Beyn's method in the following way. Define the matrices $V, W \in \mathbb{C}^{m \times n(\Gamma)}$ as follows:

$$\begin{aligned} V &= \begin{bmatrix} v_1 & \cdots & v_{n(\Gamma)} \end{bmatrix}, \\ W &= \begin{bmatrix} w_1 & \cdots & w_{n(\Gamma)} \end{bmatrix}. \end{aligned}$$

Assume that $n(\Gamma)$ is not larger than the system dimension m . In large-scale problems we actually expect to have $n(\Gamma) \ll m$. Assume also that $\text{rank}(V) = \text{rank}(W) = n(\Gamma)$, which is the case in typical applications. In case these assumptions are not satisfied, we refer to Section 3.

75 Choose $q \in \mathbb{N}$ such that $n(\Gamma) \leq q \leq m$ and choose the matrix $\hat{V} \in \mathbb{C}^{m \times q}$ such that $W^H \hat{V} \in \mathbb{C}^{n(\Gamma) \times q}$ has rank $n(\Gamma)$.

Define the matrices $S_0, S_1 \in \mathbb{C}^{m \times q}$ as follows:

$$\begin{aligned} S_0 &= \frac{1}{2\pi i} \int_{\Gamma} T(z)^{-1} \hat{V} dz, \\ S_1 &= \frac{1}{2\pi i} \int_{\Gamma} z T(z)^{-1} \hat{V} dz. \end{aligned}$$

Then

$$\begin{aligned} S_0 &= VW^H \hat{V}, \\ S_1 &= V\Lambda W^H \hat{V}, \end{aligned}$$

where the matrix $\Lambda \in \mathbb{C}^{n(\Gamma) \times n(\Gamma)}$ is defined as

$$\Lambda = \text{diag}(\lambda_1, \dots, \lambda_{n(\Gamma)}).$$

The remaining problem is the following: how to extract the eigenvalue and eigenvector information, Λ and V , from the computed matrices S_0 and S_1 ? Beyn's method [17] is based on the singular value decomposition of S_0 . Let

$$S_0 = V_0 \Sigma_0 W_0^H$$

where

$$V_0 \in \mathbb{C}^{m \times n(\Gamma)}, \quad V_0^H V_0 = I,$$

$$W_0 \in \mathbb{C}^{q \times n(\Gamma)}, \quad W_0^H W_0 = I,$$

$$\Sigma_0 = \text{diag}(\sigma_1, \dots, \sigma_{n(\Gamma)}).$$

Beyn has shown that

$$V_0^H S_1 W_0 \Sigma_0^{-1} = Q \Lambda Q^{-1}$$

where $Q = V_0^H V$. It follows that $V_0^H S_1 W_0 \Sigma_0^{-1}$ is diagonalizable. Its eigen-
 80 values are the eigenvalues of T inside the contour and its eigenvectors lead to
 the corresponding eigenvectors of T . In Section 5, we are going to present an
 alternative way of extracting the eigenvalue and eigenvector information based
 on the canonical polyadic decomposition of the tensor consisting of the two
 slices S_0 and S_1 . However, we are first going to investigate the influence of
 85 approximating the contour integrals by numerical integration.

3. Numerical integration and filter functions

For several problems, e.g., when $n(\Gamma) > m$, knowing only S_0 and S_1 is
 not sufficient to retrieve the eigenvalues and eigenvectors. Therefore, we are
 also considering the higher order moments. Define the matrices $S_p \in \mathbb{C}^{m \times q}$ as
 follows:

$$S_p = \frac{1}{2\pi i} \int_{\Gamma} z^p T(z)^{-1} \hat{V} dz, \quad p = 0, 1, 2, \dots$$

We approximate S_p by a N -point quadrature formula with nodes z_j and corres-
 ponding weights ω_j for $j = 0, 1, 2, \dots, N-1$.

$$S_p \approx \tilde{S}_p = \sum_{j=0}^{N-1} \omega_j z_j^p T(z_j)^{-1} \hat{V}, \quad p = 0, 1, 2, \dots \quad (2)$$

Note that for large-scale problems, that is with m large, most of the computa-
 tional effort comes from computing the vectors $T(z_j)^{-1} \hat{V}$ for $j = 0, 1, \dots, N-1$.
 One of the great advantages of contour integration methods is that this compu-
 90 tation can be performed in parallel for the different values of z_j by computing
 the solution of the corresponding linear system $T(z_j)X_j = \hat{V}$.

Using Keldysh' theorem, we can rewrite (2) as

$$\tilde{S}_p = \sum_{k=1}^{n(\mathcal{C})} v_k w_k^H \hat{V} \sum_{j=0}^{N-1} \frac{\omega_j z_j^p}{z_j - \lambda_k} + \sum_{j=0}^{N-1} \omega_j z_j^p R(z_j) \hat{V}. \quad (3)$$

The function $b_p : \mathbb{C} \rightarrow \mathbb{C}$ defined as

$$b_p(z) = \sum_{j=0}^{N-1} \frac{\omega_j z_j^p}{z_j - z}, \quad p = 0, 1, 2, \dots$$

is called the filter function (of order p) corresponding to the quadrature formula.

Recently there is a lot of interest studying and designing such filter functions.

We refer the interested reader to [5, 3, 20]. Using this definition for the filter

function, we can write \tilde{S}_p from (3) as

$$\tilde{S}_p = \sum_{k=1}^{n(\mathcal{C})} v_k w_k^H \hat{V} b_p(\lambda_k) + \sum_{j=0}^{N-1} \omega_j z_j^p R(z_j) \hat{V}.$$

If Γ is the unit circle, then the trapezoidal rule can be used as quadrature formula. In this case the nodes are given by

$$z_j = e^{i 2\pi j / N}$$

whereas the weights are equal to

$$\omega_j = \frac{z_j}{N}$$

for $j = 0, 1, 2, \dots, N-1$. It follows that

$$b_0(z) = \sum_{j=0}^{N-1} \frac{\omega_j}{z_j - z} = \frac{1}{N} \sum_{j=0}^{N-1} \frac{z_j}{z_j - z} = \frac{1}{1 - z^N}$$

and

$$b_p(z) = \sum_{j=0}^{N-1} \frac{\omega_j z_j^p}{z_j - z} = \frac{1}{N} \sum_{j=0}^{N-1} \frac{z_j^{p+1}}{z_j - z} = \frac{z^p}{1 - z^N}$$

for $p = 1, 2, \dots$. Therefore we may conclude that

$$b_p(z) = z^p b_0(z), \quad p = 0, 1, 2, \dots$$

in case of the unit circle and the trapezoidal rule.

4. Interpretation of Beyn's algorithm based on the filter function

We will limit our study to the unit circle case. A similar investigation can be carried out for other contours. Suppose that Γ is the unit circle. Then the ideal filter

$$\frac{1}{2\pi i} \int_{\Gamma} \frac{1}{z - \lambda} dz = \begin{cases} 1 & |\lambda| < 1 \\ 0 & |\lambda| > 1 \end{cases}$$

is approximated by

$$b_0(\lambda) = \frac{1}{1 - \lambda^N}.$$

Let $\epsilon > 0$ be small, $\epsilon \ll 1$. Then the ϵ -level curve of $b_0(\lambda)$, that is

$$\{ \lambda \in \mathbb{C} : |b_0(\lambda)| = \epsilon \}$$

is approximately a circle with its center at the origin and radius $\rho_{\epsilon, N} = \epsilon^{-1/N}$:

$$\left| \frac{1}{1 - \lambda^N} \right| = \epsilon \quad \Rightarrow \quad |\lambda| \approx \epsilon^{-1/N}.$$

Hence, for a fixed ϵ , doubling the number of quadrature nodes N squares this radius,

$$\rho_{\epsilon, 2N} = \sqrt{\rho_{\epsilon, N}}$$

and, for fixed N ,

$$2\rho_{\epsilon, N} = \rho_{\frac{\epsilon}{2^N}, N}.$$

Let $\delta \approx 10^{-16}$ denote the machine precision (double precision). Then the
95 following table indicates how $\rho_{\delta, N}$ decreases as N increases as a power of 2.

N	$\rho_{\delta, N}$
4	9741.98
8	98.70
16	9.93
32	3.15
64	1.78
128	1.33
256	1.15

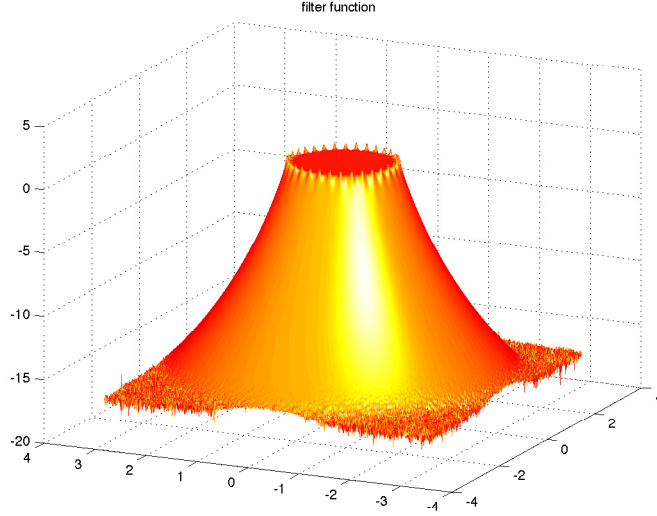


Figure 1: Plot of the log of the modulus of the filter function b_0

Figures 1 and 2 provide a graphical representation of the filter function b_0 respectively its contour lines for $N = 32$. Note that in this case, all eigenvalues lying outside the disk with radius $\rho_{\delta,32} \approx 3.15$ are filtered away up to the machine precision.

Define the contour Γ_ϵ as

$$\Gamma_\epsilon = \{\lambda \in \mathbb{C} : |b_0(\lambda)| = \epsilon\}.$$

Note that for small values of ϵ , this contour is almost a circle with radius $\rho_{\delta,N}$. Assume that Γ_ϵ is inside \mathcal{C} (defined in Theorem 1), then we can split the set of eigenvalues into two subsets: the eigenvalues located inside Γ_ϵ and those located outside or on Γ_ϵ . Let us assume that the eigenvalues located inside Γ_ϵ ,

$$\lambda_1, \dots, \lambda_{n(\Gamma_\epsilon)},$$

are ordered such that

$$|b_0(\lambda_1)| \geq \dots \geq |b_0(\lambda_{n(\Gamma_\epsilon)})|$$

and let us assume that the eigenvalues located outside Γ_ϵ ,

$$\lambda_{n(\Gamma_\epsilon)+1}, \dots, \lambda_{n(\mathcal{C})},$$

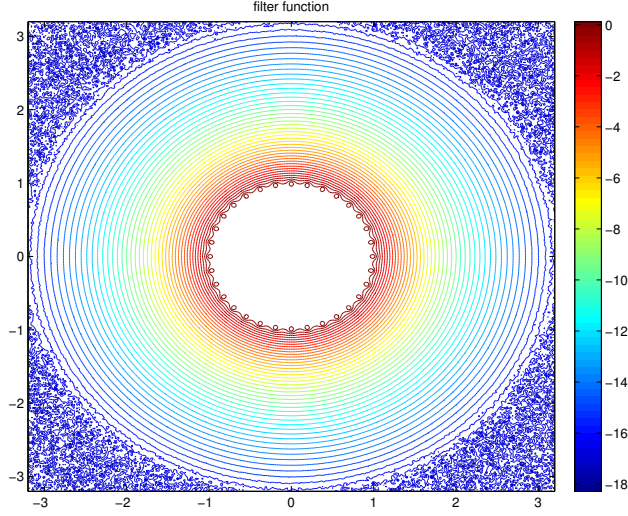


Figure 2: Contour plot of the log of the modulus of the filter function b_0

are ordered such that

$$|b_0(\lambda_{n(\Gamma_\epsilon)+1})| \geq \dots \geq |b_0(\lambda_{n(\mathcal{C})})|.$$

It follows that

$$\begin{aligned} \tilde{S}_p &= \sum_{k=1}^{n(\Gamma_\epsilon)} v_k w_k^H \hat{V} b_p(\lambda_k) \\ &\quad + \sum_{k=n(\Gamma_\epsilon)+1}^{n(\mathcal{C})} v_k w_k^H \hat{V} b_p(\lambda_k) + \sum_{j=0}^{N-1} \omega_j z_j^p R(z_j) \hat{V}. \end{aligned}$$

We know that

$$|b_0(\lambda_k)| \leq |b_0(\lambda_{n(\Gamma_\epsilon)+1})| \lesssim \epsilon$$

for $k = n(\Gamma_\epsilon) + 1, \dots, n(\mathcal{C})$.

The convergence radius r of $R(z)$ satisfies

$$r \gtrsim \rho_{\epsilon, N}.$$

The exact value of r is determined by the minimum of the following two values: the smallest modulus of the eigenvalues outside \mathcal{C} and the smallest modulus of

105 the points for which T is not analytic. Note again that for polynomial and linear eigenvalue problems with nonsingular highest degree coefficient the function R is equal to zero.

Suppose we want to approximate the eigenvalues inside Γ_ϵ . So, we consider the second and third sum in the expression for \tilde{S}_p as error terms. Let us examine the norm of these terms. Define $\hat{\lambda}$ as

$$\hat{\lambda} = \lambda_{n(\Gamma_\epsilon)+1}.$$

Then the norm of the dominant component of the second sum is equal to

$$c_1 |b_p(\hat{\lambda})| = c_1 \left| \frac{\hat{\lambda}^p}{1 - \hat{\lambda}^N} \right| \approx c_1 |\hat{\lambda}|^{p-N} \leq c_1 \epsilon^{1-\frac{p}{N}}$$

assuming $p \leq N$. Based on Cauchy's estimate, one can show that the norm of the third sum behaves as

$$c_2 r^{p-N} \lesssim c_2 \epsilon^{1-\frac{p}{N}}$$

because

$$r \gtrsim \rho_{\epsilon,N} = \epsilon^{-1/N}.$$

In summary, we obtain that

$$\tilde{S}_p = \sum_{k=1}^{n(\Gamma_\epsilon)} v_k w_k^H \hat{V} b_p(\lambda_k) + \Delta_1 + \Delta_2 \quad (4)$$

with

$$\|\Delta_1\| \approx c_1 |\hat{\lambda}|^{p-N} \leq c_1 \epsilon^{1-\frac{p}{N}}$$

and

$$\|\Delta_2\| \approx c_2 r^{p-N} \lesssim c_2 \epsilon^{1-\frac{p}{N}}.$$

The terms Δ_1 and Δ_2 can be considered as error terms contaminating the “true” value

$$\bar{S}_p = \sum_{k=1}^{n(\Gamma_\epsilon)} v_k w_k^H \hat{V} b_p(\lambda_k). \quad (5)$$

Note that to keep these error terms small, it is best to keep p small compared to N .

Besides the error terms Δ_1 and Δ_2 of order $\epsilon^{1-\frac{p}{N}}$, not all eigenvalues are equally represented in the sum \tilde{S}_p . E.g., for S_0 , the more the eigenvalues are away from the unit disk, that is the smaller $|b_0(\lambda_k)|$, the more the term corresponding to this eigenvalue is filtered away in the sum S_0 . The different strategies to solve the nonlinear eigenvalue problem of Section 6 indicate that, under certain conditions, one can retrieve v_k (2-norm 1) and λ_k with relative residuals of magnitude

$$c_3 \frac{\epsilon}{b_0(\lambda_k)}.$$

110 A more detailed analysis is necessary.

5. Robust extraction of the eigenvalues λ_k and the eigenvectors v_k from the computed moments \tilde{S}_p

We now present an alternative for Beyn's approach. We recall from (4) that

$$\tilde{S}_p = \sum_{k=1}^{n(\Gamma_\epsilon)} v_k w_k^H \hat{V} b_p(\lambda_k) + \Delta_1 + \Delta_2$$

with $\|\Delta_1\| + \|\Delta_2\|$ having size $c_3 \epsilon^{1-\frac{p}{N}}$ and that

$$b_p(\lambda_k) = \lambda_k^p b_0(\lambda_k)$$

in case of the unit circle and the trapezoidal rule. Up to the error terms Δ_1 and Δ_2 , it follows that

$$\tilde{S}_p \approx \sum_{k=1}^{n(\Gamma_\epsilon)} v_k w_k^H \hat{V} \lambda_k^p b_0(\lambda_k) = V \Lambda^p \hat{W}^H$$

where

$$\begin{aligned} V &= \begin{bmatrix} v_1 & \cdots & v_{n(\Gamma_\epsilon)} \end{bmatrix}, \\ \Lambda &= \text{diag}(\lambda_1, \dots, \lambda_{n(\Gamma_\epsilon)}), \\ W &= \begin{bmatrix} w_1 & \cdots & w_{n(\Gamma_\epsilon)} \end{bmatrix}, \\ \hat{W}^H &= \text{diag}(b_0(\lambda_1), \dots, b_0(\lambda_{n(\Gamma_\epsilon)})) W^H \hat{V}. \end{aligned}$$

Define H and $H^<$ as

$$H = \begin{bmatrix} \hat{S}_0 \end{bmatrix} \approx V \hat{W}^H$$

and

$$H^< = \begin{bmatrix} \hat{S}_1 \end{bmatrix} \approx V \Lambda \hat{W}^H.$$

¹¹⁵ In general, let

$$\begin{aligned} H &= \begin{bmatrix} \hat{S}_0 & \hat{S}_1 & \cdots & \hat{S}_\nu \\ \hat{S}_1 & & \ddots & \vdots \\ \vdots & \ddots & & \vdots \\ \hat{S}_\mu & \cdots & \cdots & \hat{S}_{\mu+\nu} \end{bmatrix} \\ &\approx \begin{bmatrix} V \\ V\Lambda \\ \vdots \\ V\Lambda^\mu \end{bmatrix} \begin{bmatrix} \hat{W}^H & \Lambda \hat{W}^H & \cdots & \Lambda^\nu \hat{W}^H \end{bmatrix} \end{aligned}$$

and let

$$\begin{aligned} H^< &= \begin{bmatrix} \hat{S}_1 & \hat{S}_2 & \cdots & \hat{S}_{\nu+1} \\ \hat{S}_2 & & \ddots & \vdots \\ \vdots & \ddots & & \vdots \\ \hat{S}_{\mu+1} & \cdots & \cdots & \hat{S}_{\mu+\nu+1} \end{bmatrix} \\ &\approx \begin{bmatrix} V \\ V\Lambda \\ \vdots \\ V\Lambda^\mu \end{bmatrix} \Lambda \begin{bmatrix} \hat{W}^H & \Lambda \hat{W}^H & \cdots & \Lambda^\nu \hat{W}^H \end{bmatrix} \end{aligned}$$

be defined as block Hankel matrices of block size $(\mu + 1) \times (\nu + 1)$.

Therefore, the canonical polyadic decomposition [21] of the tensor consisting of the two slices H and $H^<$ is given by

$$\sum_{k=1}^{n(\Gamma_e)} \begin{bmatrix} v_k \\ v_k \lambda_k \\ \vdots \\ v_k \lambda_k^\mu \end{bmatrix} \odot \begin{bmatrix} \overline{\hat{w}_k} \\ \lambda_k \overline{\hat{w}_k} \\ \vdots \\ \lambda_k^\nu \overline{\hat{w}_k} \end{bmatrix} \odot \begin{bmatrix} 1 \\ \lambda_k \end{bmatrix}$$

where \hat{w}_k denotes the k th column vector of \hat{W} for $k = 1, \dots, n(\Gamma_\epsilon)$ and \odot denotes the outer product.

At the other end of the spectrum of all structured tensors that one can build from the computed moments \hat{S}_p , $p = 1, 2, \dots, P$, we can consider the tensor whose slices are these moments. In this case the canonical polyadic decomposition is of the form

$$\sum_{k=1}^{n(\Gamma_\epsilon)} v_k \odot \overline{\hat{w}_k} \odot \begin{bmatrix} 1 \\ \lambda_k \\ \vdots \\ \lambda_k^P \end{bmatrix}.$$

120 All other possible “Hankel” tensors based on the moments can be considered too. Using a tensor approach instead of a matrix approach can lead to a more robust algorithm [22, 23]. Tensorlab [24] provides a robust algorithm for computing this canonical polyadic decomposition. In our numerical experiments, the tensor variant has always computed solutions that are at least as accurate as via Beyn’s
125 method. However, further analysis is necessary.

We are now able to formulate the skeleton of our algorithm for solving non-linear eigenvalue problems via contour integrals.

1. Choose a filter function, that is a quadrature formula,
and where to apply it in the complex plane based on $T(z)$ and the domain
130 in which the requested eigenvalues are lying.
2. Choose $\hat{V} \in \mathbb{C}^{m \times q}$ of rank q .
3. Compute the moments \hat{S}_p , $p = 1, \dots, P$ via contour integration based on the quadrature formula.
4. Compute the canonical polyadic decomposition of a Hankel tensor constructed from the computed moments \hat{S}_p .
135
5. The first and third factor matrix lead to the approximate eigenvectors and corresponding eigenvalues.

6. Numerical experiment

We have tested our algorithm on several linear, polynomial and nonlinear eigenvalue problems. Instead of presenting an overview of these experiments, we consider only one example and choose different scenarios to solve this nonlinear eigenvalue problem. We consider the gun problem from the collection of nonlinear eigenvalue problems established by Betcke, Higham, Mehrmann, Schroeder and Tisseur [25]. This problem is related to a model of a radio-frequency gun cavity. The problem size is equal to $m = 9956$ whereas the function T has the following form:

$$T(z) = \begin{bmatrix} K & M & iW_1 & iW_2 \end{bmatrix} \begin{bmatrix} 1 \\ -z \\ \sqrt{z} \\ \sqrt{z - \alpha} \end{bmatrix}$$

where K , M , W_1 and W_2 are sparse matrices, and $\alpha = (108.8774)^2$.

140 We would like to approximate all the eigenvalues located inside the circle that is symmetric with respect to the real axis, and that intersects the real axis at $\alpha = (108.8774)^2$ and $\beta = 340^2$.

We choose Γ as the unit circle and $N = 32$.

What is a good choice for the center μ and the radius ρ such that we can
145 apply the theory to $T(\mu + \rho z)$?

Let us consider different strategies.

Scenario 1. In this case, we choose the parameters μ and ρ such that $z = -3.1$ is mapped to the branch point α and such that the point $z = 2$ is mapped to the point β .

z	$ b_0(z) $	$\mu + \rho z$
-3.1	$\approx 10^{-16}$	α
2	$\approx 10^{-10}$	β

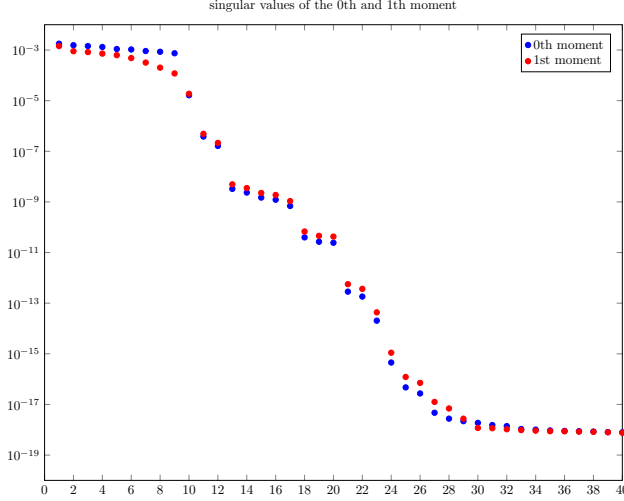


Figure 3: Singular values of \tilde{S}_0 and \tilde{S}_1

Because the convergence radius r of $R(z)$ is equal to 3.1 and

$$|b_0(r)| \approx 10^{-16}$$

the $R(z)$ part does not influence (up to machine precision) the computation of \tilde{S}_0 and \tilde{S}_1 . In other words,

$$\begin{aligned}\tilde{S}_0 &= \sum_{k=1}^{n(\mathcal{C})} v_k w_k^H \hat{V} b_0(\lambda_k), \\ \tilde{S}_1 &= \sum_{k=1}^{n(\mathcal{C})} v_k w_k^H \hat{V} \lambda_k b_0(\lambda_k)\end{aligned}$$

up to machine precision. This is indicated by the singular values of \tilde{S}_0 and \tilde{S}_1 shown in Figure 3. One can clearly observe the different $|b_0(\lambda_k)|$ levels / sizes.

155 Note that the difference between the blue and red dots gives an indication of the modulus of the corresponding eigenvalue.

By computing the canonical polyadic tensor decomposition with 23 terms, we obtain an approximation of 23 eigenvalues.

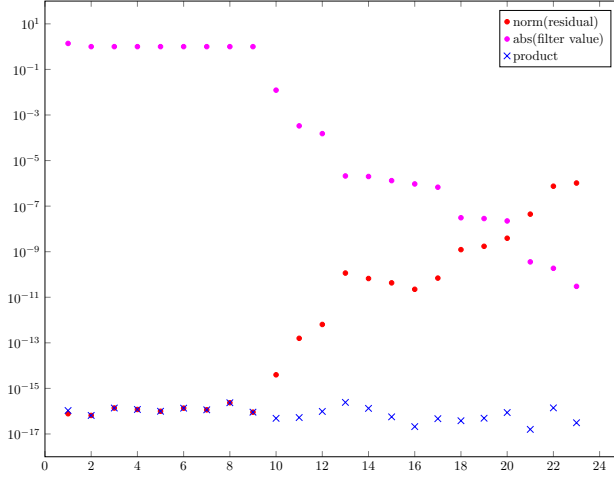


Figure 4: Relative residuals for the computed eigenvalues

The relative residual for each of the computed eigenvalues $\tilde{\lambda}_k$ with corresponding eigenvector \tilde{v}_k is given by

$$\epsilon_k = \frac{\|T(\tilde{\lambda}_k)\tilde{v}_k\|_1}{\|T(\tilde{\lambda}_k)\|_1\|\tilde{v}_k\|_1}.$$

Figure 4 plots these relative residuals. Because the accuracy of $\tilde{\lambda}_k$ depends on the extraction of the corresponding information from \tilde{S}_0 and \tilde{S}_1 , this accuracy is limited to $|b_0(\tilde{\lambda}_k)|$. Hence, we expect that

$$\epsilon_k |b_0(\tilde{\lambda}_k)| \approx |b_0(r)|.$$

Let us check that

$$\|R_\alpha\| \approx \gamma \left(\frac{1}{r}\right)^\alpha$$

where $r = 3.1$ is the expected convergence radius of $R(z) = \sum_{\alpha \geq 0} R_\alpha z^\alpha$. One can approximate the norms $\|R_\alpha\|$ by using the Fast Fourier Transform. We are not giving the details here. The results shown in Figure 5 imply that the computed convergence radius $\tilde{r} \approx 2.55$. We miss an eigenvalue λ^* having the following properties;

$$|b_0(\lambda^*)| \approx 10^{-13}$$

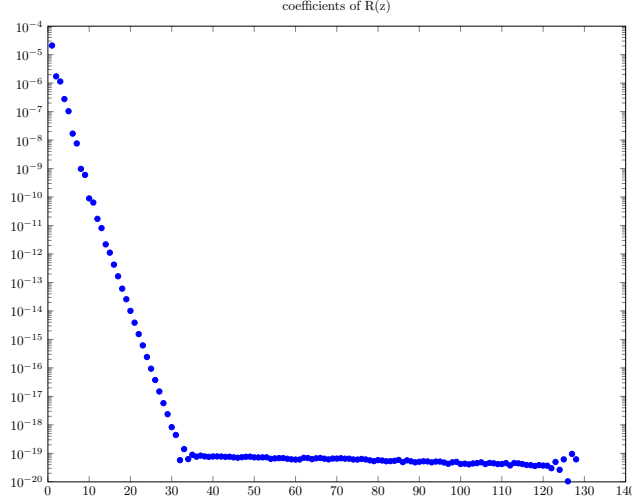


Figure 5: Coefficients of $R(z)$

and

$$|\lambda^*| \approx 2.58 \approx \tilde{r}.$$

The two filled red dots in Figure 6 indicate α and β . The red circle represents the
 160 eigenvalue λ^* that is missed, while the other blue circles represent the computed
 eigenvalues $\tilde{\lambda}_k$. Let us try another scenario.

Scenario 2. In this case, we choose the parameters μ and ρ such that $z = -2$ is
 mapped to the branch point α and such that the point $z = 2$ is mapped to the
 point β .

z	$ b_0(z) $	$\mu + \rho z$
-2	$\approx 10^{-10}$	α
2	$\approx 10^{-10}$	β

The convergence radius r of $R(z)$ now equals $r = 2$ and

$$|b_0(r)| \approx 10^{-10}.$$

In the computation of the moments \tilde{S}_0 and \tilde{S}_1 the $R(z)$ term now comes in
 with size of order 10^{-10} . Figure 7 shows the singular values of \tilde{S}_0 and \tilde{S}_1 .

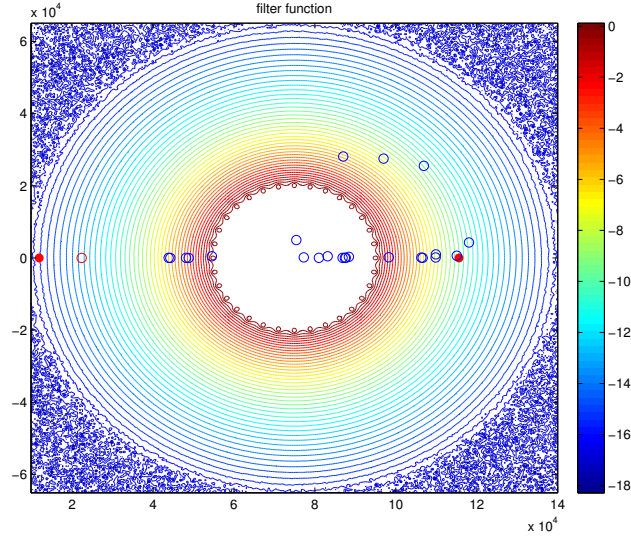


Figure 6: Eigenvalues

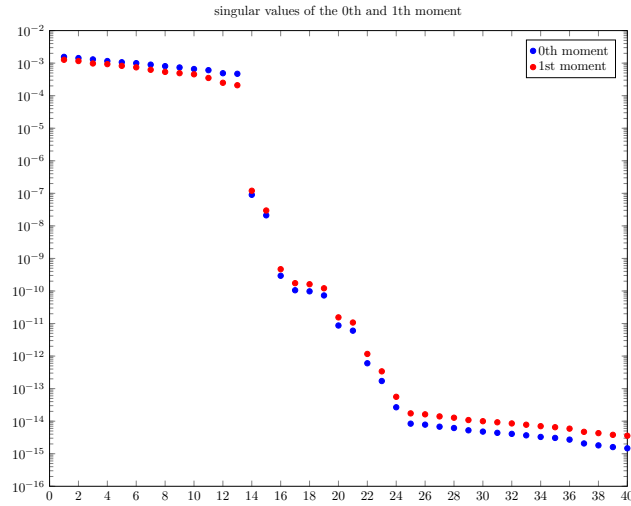


Figure 7: Singular values of \tilde{S}_0 and \tilde{S}_1

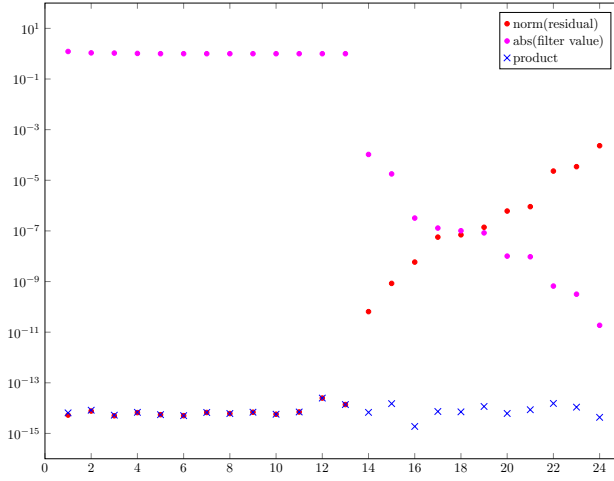


Figure 8: Relative residuals for the computed eigenvalues

By computing the canonical polyadic tensor decomposition with 24 terms, we obtain an approximation of 24 eigenvalues. Figure 8 plots the relative residuals.

170

Let us check that

$$\|R_\alpha\| \approx \gamma \left(\frac{1}{r}\right)^\alpha$$

where $r = 2$ is the convergence radius of $R(z) = \sum_{\alpha \geq 0} R_\alpha z^\alpha$. The results shown in Figure 9 imply that $\tilde{r} \approx 2$. Note that R_0 is of the order 10^{-5} which explains why we get smaller values of the relative residual in Figure 8. Comparing Figures 6 and 10 one can observe that we have found the missing eigenvalue.

175 However, the accuracy of this eigenvalue will be small. To derive this eigenvalue with high accuracy, we consider the following strategy.

Scenario 3. Let us consider a filter function having the following property: the eigenvalues that are lying in the neighbourhood of a branch point have a larger value of the modulus of the filter function compared to the “classical” filter functions.

180

Consider, for example, the filter function developed by Hale, Higham and Trefethen [26] to handle the branch point $z = 0$ of \sqrt{z} . A plot of the modulus

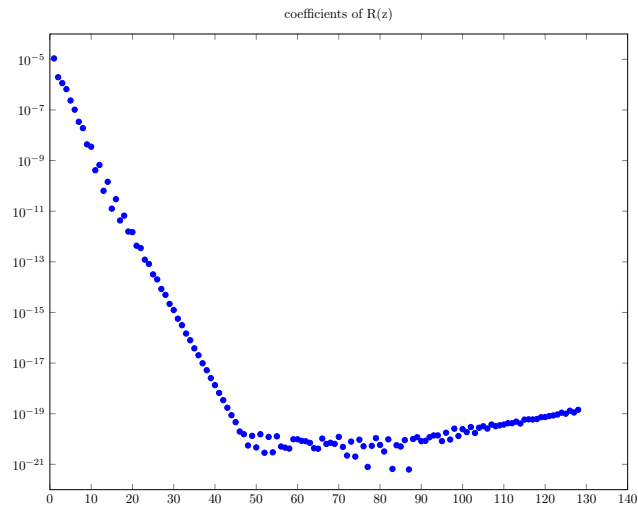


Figure 9: Coefficients of $R(z)$

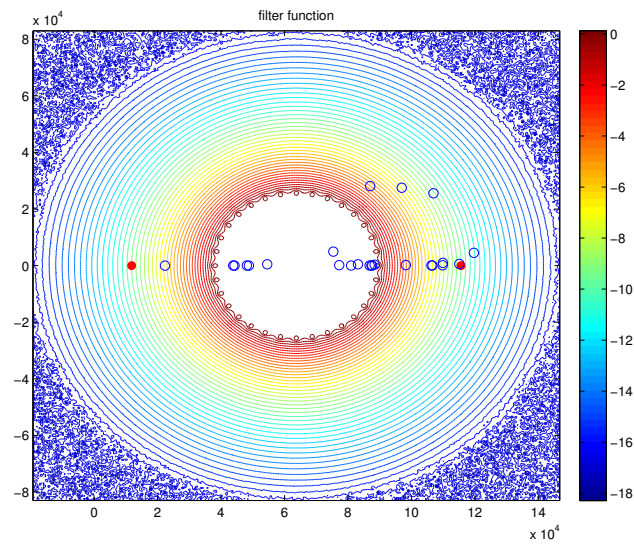


Figure 10: Eigenvalues

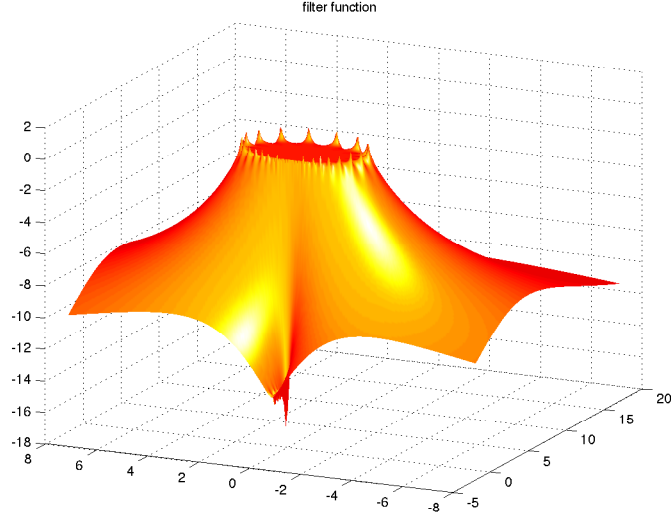


Figure 11: Plot of the log of the modulus of the filter function developed by Hale, Higham and Trefethen

of this filter function is shown in Figure 11.

The transformation parameters μ and ρ are determined such that

$$\begin{array}{rcl} z & |b_0(z)| & \mu + \rho z \\ -0.05 & \approx 10^{-16} & \alpha \\ 10 & \approx 10^{-7} & \beta \end{array}$$

The singular values of \tilde{S}_0 and \tilde{S}_1 are shown in Figure 12.

By computing the canonical polyadic tensor decomposition with 40 terms, we obtain an approximation of 40 eigenvalues. Figure 13 plots the relative residuals. Note that this particular filter function allows to compute the eigenvalues in the neighbourhood of the branch point α with a higher accuracy compared to Scenario 2.

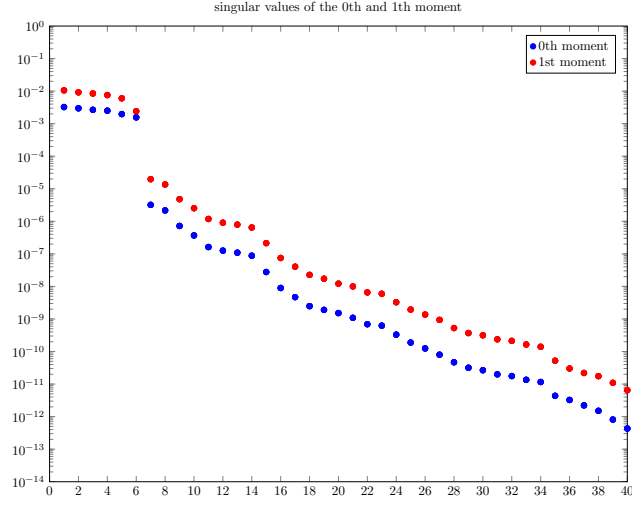


Figure 12: Singular values of \tilde{S}_0 and \tilde{S}_1

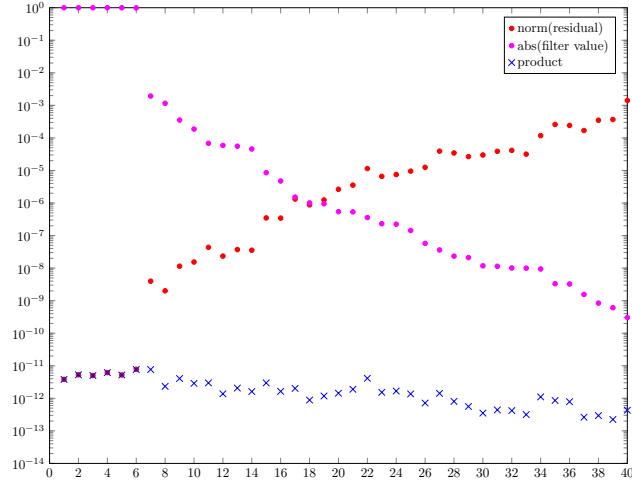


Figure 13: Relative residuals for the computed eigenvalues

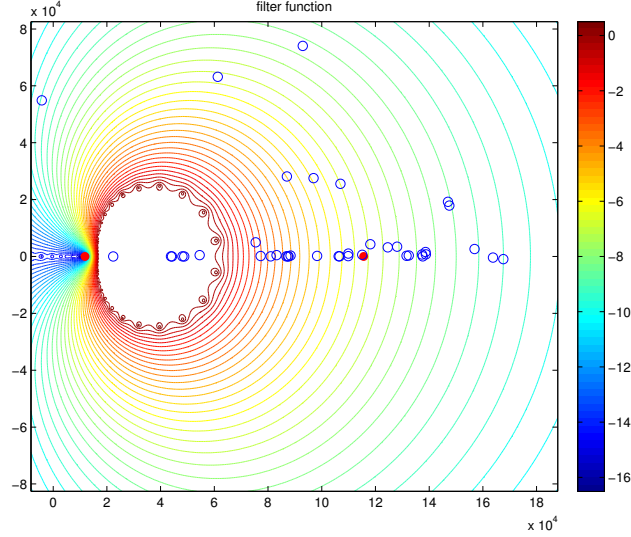


Figure 14: Eigenvalues

7. Conclusion

In this paper, we have shown that the accuracy of the computed eigenvalues and eigenvectors using contour integral “measurements” of the resolvent not only depends on the eigenvalues inside the contour Γ but also on the eigenvalue configuration outside the contour Γ where they are not completely filtered away by the filter function. In contrast to linear and polynomial eigenvalue problems with nonsingular highest degree coefficient, the convergence radius of the R -term in the expression for the resolvent plays also a crucial role in assessing the accuracy of the computed eigenvalues. This convergence radius is determined by the points z in the neighbourhood of the contour Γ for which $T(z)$ is not analytic. In the third scenario we have used a rational filter function that allows to approximate eigenvalues very accurately in the neighbourhood of a branch point of $T(z)$. In future research, we would like to develop an algorithm that designs good rational filters given the region in which one wants to approximate the eigenvalues accurately as well as the points in which $T(z)$ exhibits non analytic behaviour.

References

- [1] P. Kravanja, M. Van Barel, Computing the zeros of analytic functions, Vol. 1727 of Lecture Notes in Mathematics, Springer, 2000.
- [2] L. M. Delves, J. N. Lyness, A numerical method for locating the zeros of an analytic function, *Mathematics of Computation* 21 (1967) 543–560.
- [3] A. P. Austin, P. Kravanja, L. N. Trefethen, Numerical algorithms based on analytic function values at roots of unity, *SIAM Journal on Numerical Analysis* 52 (2) (2014) 1795–1821.
- [4] T. Sakurai, H. Sugiura, A projection method for generalized eigenvalue problems using numerical integration, *Journal of Computational and Applied Mathematics* 159 (2003) 119–128.
- [5] T. Ikegami, T. Sakurai, U. Nagashima, A filter diagonalization for generalized eigenvalue problems based on the Sakurai-Sugiura projection method, *Journal of Computational and Applied Mathematics* 233 (2010) 1927–1936.
- [6] H. Ohno, Y. Kuramashi, T. Sakurai, H. Tadano, A quadrature-based eigensolver with a Krylov subspace method for shifted linear systems for Hermitian eigenproblems in lattice QCD, *JSIAM Letters* 2 (2010) 115–118.
- [7] T. Ikegami, T. Sakurai, Contour integral eigensolver for non-Hermitian systems: A Rayleigh-Ritz-type approach, *Taiwanese J. Math.* 14 (3A) (2010) 825–837.
- [8] T. Sakurai, H. Tadano, CIRR: A Rayleigh-Ritz type method with contour integral for generalized eigenvalue problems, *Hokkaido Math. J.* 36 (2007) 745–757.
- [9] E. Polizzi, Density-matrix-based algorithm for solving eigenvalue problems, *Physics Review B* 79.

- [10] L. Krämer, E. Di Napoli, M. Galgon, B. Lang, P. Bientinese, Dissecting the FEAST algorithm for generalized eigenproblems, *Journal of Computational and Applied Mathematics* 244 (2013) 1–9.
- [11] P. T. P. Tang, E. Polizzi, FEAST as a subspace iteration eigensolver accelerated by approximate spectral projection, to appear in *SIAM J. Matrix Anal. Appl.*
- [12] S. E. Laux, Solving complex band structure problems with the FEAST eigenvalue algorithm, *Physics Review B* 79.
- [13] S. Güttel, E. Polizzi, P. T. P. Tang, G. Viaud, Zolotarev quadrature rules and load balancing for the FEAST eigensolver, MIMS EPrint 2014.39, Manchester Institute for Mathematical Sciences, The University of Manchester (Jul. 2014).
- [14] S. Yokota, T. Sakurai, A projection method for nonlinear eigenvalue problems using contour integrals, *JSIAM Letters* 5 (2013) 41–44.
- [15] J. Asakura, T. Sakurai, H. Tadano, T. Ikegami, K. Kimura, A numerical method for nonlinear eigenvalue problems using contour integrals, *JSIAM Letters* 1 (2009) 52–55.
- [16] J. Asakura, T. Sakurai, H. Tadano, T. Ikegami, K. Kimura, A numerical method for polynomial eigenvalue problems using contour integrals, *Japan J. Indust. Appl. Math.* 27 (2010) 73–90.
- [17] W.-J. Beyn, An integral method for solving nonlinear eigenvalue problems, *Linear Algebra and its Applications* 436 (2012) 3839–3863.
- [18] M. V. Keldysh, On the characteristic values and characteristic functions of certain classes of non-self-adjoint equations, *Doklady Akademii Nauk* 77 (1951) 11–14.
- [19] M. V. Keldysh, The completeness of eigenfunctions of certain classes of nonselfadjoint linear operators, *Uspehi Mat. Nauk* 26 (4160) (1971) 15–41.

- 260 [20] A. P. Austin, L. N. Trefethen, Computing eigenvalues of real symmetric matrices with rational filters in real arithmetic, SIAM Journal on Scientific Computing, submitted.
URL https://people.maths.ox.ac.uk/austin/pubs/AT2014_Preprint.pdf
- 265 [21] T. G. Kolda, B. W. Bader, Tensor decompositions and applications, SIAM Review 51 (3) (2009) 455–500.
- [22] J. M. Papy, L. De Lathauwer, S. Van Huffel, Exponential data fitting using multilinear algebra: the decimative case, Journal of Chemometrics 23 (7-8) (2009) 341–351.
- 270 [23] J. M. Papy, L. De Lathauwer, S. Van Huffel, Exponential data fitting using multilinear algebra: the single-channel and the multi-channel case, Numerical Linear Algebra with Applications 12 (8) (2005) 809–826.
- [24] L. Sorber, M. Van Barel, L. De Lathauwer, Tensorlab v2.0 (Jan. 2014).
URL <http://www.tensorlab.net>
- 275 [25] T. Betcke, N. J. Higham, V. Mehrmann, C. Schröder, F. Tisseur, NLEVP: A collection of nonlinear eigenvalue problems, MIMS EPrint 2011.116, Manchester Institute for Mathematical Sciences, The University of Manchester (Dec. 2011).
- 280 [26] N. Hale, N. J. Higham, L. N. Trefethen, Computing A^α , $\log(A)$, and related matrix functions by contour integrals, SIAM Journal on Numerical Analysis 46 (5) (2008) 2505–2523.